DOCUMENT RESUME

ED 256 342 IR 011 662

AUTHOR Torardi, Mary Montag
TITLE The Development of a Computer Literacy Assessment Instrument.
PUB DATE Jan 85
NOTE 29p.; Paper presented at the Annual Convention of the Association for Educational Communications and Technology (Anaheim, CA, January 17-23, 1985). For entire proceedings, see IR 011 621.
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; Competence; *Computer Literacy; *Computers; *Computer Science Education; Field Tests; Higher Education; Item Analysis; Models; Skill Analysis; Student Attitudes; *Test Construction; *Test Reliability; *Test Validity

ABSTRACT

This report describes the procedures used to construct and validate the Standardized Test of Computer Literacy (STCL), a criterion-referenced instrument designed to assess students' computer literacy. Following a statement of the study problem and purpose, a description of the study methodology outlines the use of a 12-step model for development and validation of criterion-referenced tests as a guide for creating the STCL. The procedures discussed include: (1) defining the content and identifying objectives for the areas of computer attitudes, computer systems, computer applications, and computer programs; (2) gathering computer literacy competencies from the literature, a survey of 90 computer education specialists, and the Computer Anxiety Index and the Beliefs About Computers Scale; (3) development of 156 test items by 15 Iowa State University computer literacy specialists; (4) pilot testing involving empirical item analysis based on administration to 44 college students and judgmental item analysis by 9 computing instructors; and (5) collecting reliability, validity, and normative data on the final 80-item test. Results reported indicate that the STCL was both a valid and reliable measure of computer literacy. Twenty-two references are listed. (LMM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED256342

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

THE DEVELOPMENT

OF A

COMPUTER LITERACY ASSESSMENT INSTRUMENT

BY

MARY MONTAG TORARDI

IR011662

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Michael Simonson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Statement of the Problem

Educators have recently been faced with the task of meeting the demands of society to educate students so that they become computer literate. In 1983, the "Nation at Risk" report of the National Commission on Excellence in Education recommended that computer science be required of all secondary school students (National Commission on Excellence in Education, 1983). An extensive review of the literature revealed that many public school districts were developing computer literacy programs and expecting students to become computer literate. In addition, many higher educational institutions had instigated computer literacy requirements for graduation. However, there were no known standards for determining what aspects of computer literacy should be taught at various levels of education, and no effective means of measuring the level of computer literacy possessed at various levels of education. A standardized measure of computer literacy was needed to facilitate the development of computer literacy courses that were appropriate to the needs of the students. Such an instrument would be critical for providing students with instruction and experiences suited to their level of computer literacy.

Purpose of the Study

The purpose of this study was to construct a valid and reliable criterion-referenced instrument to assess the computer literacy (knowledge, skills, attitudes) of students. Three primary purposes for the Standardized Test of Computer Literacy (STCL) were identified.

1. To document the computing competencies of students and to collect normative data about various subgroups of students.
2. To justify revisions in the content of computing courses to reflect student needs.

3. To provide criteria for the placement of students in appropriate computing courses based on their level of computer literacy.

Methodology

The procedures followed in the development of the STCL were designed to contribute positively to the overall validity and reliability of the instrument. A twelve step model for the development and validation of criterion-referenced tests was used as a guide for the development of the STCL. The methodology included the following steps (Hambleton, 1980):

1. Preparation and/or selection of objectives.
2. Preparation of test domain specifications.
3. Writing test items "matched" to objectives.
4. Preliminary review of test items.
5. Determination of content validity of the test items.
   a. Involvement of content specialists.
   b. Collection and analysis of examinee response data.
6. Additional editing of test items.
7. Test assembly.
8. Setting standards for interpreting examinee test performance.
9. Test administrations to criterion groups.
10. Assessment of test score reliability and validity; compilation of test score norms.
11. Preparation of a user's manual and a technical manual.
12. Periodic collection of additional technical information.

The first step in the development of the STCL involved defining the domain of content (computer literacy) to be measured by the instrument and identifying appropriate computer literacy objectives or competencies. A review of the literature was conducted to determine a definition of computer literacy, identify computer literacy competencies, and identify any existing test items or instruments measuring computer literacy. The following definition was adopted:

> Computer literacy is operationally defined as an understanding of computer characteristics, capabilities, and applications, as well as an ability to implement this knowledge in the skillful, productive use of computer applications suitable to individual roles in society. The knowledge and skills of computer literacy as defined above have been divided into four major categories; computer attitudes, computer systems, computer applications, and computer programming. These four sections are defined below.

1. Computer Attitudes refer to an individual's feelings about the personal and societal use of computers in appropriate ways. Positive attitudes include an anxiety free willingness or desire to use the computer, confidence in one's abilities to use the computer, and computer responsibility.

2. Computer Systems refer to the appropriate, knowledgeable use of equipment (hardware) and programs (software) necessary for computer applications. This requires understanding and abilities in the following areas: computer functions, computer hardware, computer software, computer systems configuration, computer terminology, historical development, and the operation of computers.

3. Computer Applications refer to the ability to responsibly evaluate, select, and implement a variety of practical computer applications to do meaningful and efficient work based on an understanding of the following: general types of applications, capabilities and limitations of applications, societal impact (past, present, and future), evaluation and selection techniques, and specific applications (word processing, data base management, spreadsheet/financial management, statistical analysis, graphics, and educational applications).

4. Computer Programming refers to the ability to direct the operation of the computer through the skillful use of programming languages (high level as well as software languages). This requires an understanding of problem solving strategies, algorithms and flowcharts, languages, and programming skills.

Next, computer literacy competencies were collected from a variety of sources, such as computer literacy course outlines, curriculum guides, textbooks, and computer literacy research studies reported in the literature. "The Computer Literacy Objectives", developed as part of the Computer Literacy Study conducted by the Minnesota Educational Computing Consortium (Klassen et al., 1980), were reviewed and accepted for the initial list of computer literacy competencies.

An attempt was made to ensure a more comprehensive and valid list of competencies for the STCL domain specifications. A survey was sent to 327 computer education specialists. Each computer specialist was asked to read the purpose of the STCL and the operational definition of computer literacy, and then write two computer literacy competencies for each of the four sections of the definition. The computer specialists were selected from the following

sources:

1. The participants of the National Computer Literacy Goals for 1985 Conference (Seidel, Anderson, and Hunter, 1982).

2. The presenters at the National Educational Computing Conference of 1982 (Smith and Moum, 1982).

3. Computer Consultants listed in the 1983 Classroom Computer News Directory of Educational Computing Resources, Part III: Local and Regional Resources (Kelman, 1983).

4. The members of the International Council for Computers in Education, listed in the April 1983 issue of The Computing Teacher (Moursund, 1983).

5. Names of instructors derived from a listing of colleges and universities offering summer school computer courses in the March 1982 issue of The Computing Teacher (Moursund, 1982).

The list of computer literacy competencies obtained from approximately 90 computer specialists who responded to the survey was then combined with the list of competencies identified from the literature. A tally of each discreet competency was made by counting the number of individuals who mentioned each competency. The entire list of competencies was then reviewed by the STCL steering committee. A total of 87 competencies were selected for inclusion in the STCL domain specifications based on the following criteria (Klein and Kosecoff, 1973): transferability within the domain, widely accepted by content specialists, terminality, transferability outside the domain, and ease of scorability. The competencies were then categorized into the four sections of the definition (nine for computer attitudes, 24 for computer systems, 33 for computer applications, and 21 for computer programming).

The review of the literature revealed three valid and reliable computer-related instruments that were reviewed for possible inclusion in the STCL. "The Minnesota Computer Literacy and Awareness Assessment" (Klassen et al., 1980), The Computer Anxiety Index (CAIN) (Maurer, 1983), and the Beliefs About Computers Scale (BACS) (Ellsworth and Bowman, 1982) were reviewed to determine if any portions of the instruments were consistent with the computer literacy definition and competencies for the STCL.

It was the opinion of the steering committee that "The Minnesota Computer Literacy and Awareness Assessment" did not contain test items that were appropriate for the STCL. Most of the items on the Minnesota instrument were knowledge or comprehension level questions. One of the goals identified by the STCL steering committee was to develop an instrument that included test items that measured high level cognitive skills as identified by Bloom's Taxonomy of Educational Objectives (Bloom, Englehart, Hill, and Krathwohl, 1956). An attempt was made to develop an instrument that measured high level computer literacy skills, such as application, analysis, synthesis, and evaluation skills. The Minnesota instrument seemed to measure only knowledge and awareness aspects of computer literacy as defined for the STCL. The STCL steering committee concluded that higher level items could be written to measure computer-related skills.

The CAIN (Maurer, 1983) and the BACS (Ellsworth and Bowman, 1982) were considered to be appropriate for the computer attitudes section of the STCL. Both of these instruments consisted of Likert-type items and were shown to be valid and reliable measures of students' attitudes and anxiety towards computers. The CAIN is a twenty-six item scale with a test/retest reliability of 0.90 and an internal consistency reliability estimate of 0.94 (Maurer, 1983). The BACS is a seventeen item scale with a test/retest reliability of 0.85 and an internal consistency reliability of 0.77 (Ellsworth and Bowman, 1982).

Domain specifications for the STCL were developed based on the suggestions proposed by Popham (1980) and others (Berk, 1980; Millman, 1974; Hively et al., 1968). The Domain specifications contained a description of the purpose of the instrument, the definition of computer literacy, the list of competencies, guidelines for writing effective multiple choice items, and sample items.

The STCL domain specifications were distributed to a team of fifteen computer literacy specialists in the department of Professional Studies in Education at Iowa State University. Each specialist was asked to write multiple choice questions for ten or eleven of the competencies in the Computer Systems, Computer Applications, and Computer Programming sections of the STCL. Two questions were written for each competency in the three sections mentioned above. This produced a total of 186 test questions. The test item writers were encouraged to incorporate diagrams, illustrations, and other creative devices into appropriate questions to ensure a variety of test items at all levels of Bloom's Taxonomy of Educational Objectives (Bloom et al., 1956).

The 186 test items written by the computer specialists were reviewed, edited, and revised by the principal researcher. The entire collection of items was then divided into two sets, each containing approximately 90 questions related to the competencies plus fourteen items involving demographic information. The test items were then pilot tested to collect data for an empirical item analysis so that revisions could be made to improve the content validity of the items.

The two sets of test items were each administered to two criterion groups as recommended by Berk (1980): an instructed group of students and an uninstructed group of students. The instructed group consisted of forty-four subjects: forty-one college students enrolled in two educational computing courses and three computing instructors in the College of Education at Iowa State University. The students in these courses were considered instructed in computer literacy because a course titled Introduction to Computer Applications was a prerequisite to each course. The uninstructed group consisted of thirty-five students in education courses who had not taken a computer literacy course previously. Students in these courses who reported that they had taken a computer course were not included in the data analysis.

The data from each set of items, each section of the test, and each criterion group were analyzed by obtaining the following statistics: means, variances, standard deviations, item difficulty indices, item discrimination indices, and the frequencies of students responding to each distractor on each item. These statistics were carefully analyzed to identify items that needed to be revised or discarded to improve the content validity of the test.

In addition to the empirical item analysis, a judgmental item analysis was conducted by a panel of nine computer specialists at Iowa State University. The computer specialists were asked to examine each item and its respective competencies and make a judgment about the item-competency congruence and the technical quality of each item. A rating scale like one proposed by Rovinelli and Hambleton (1976) was used to rate how well each item measured the competency it was intended to measure (1-poor, 2-fair, 3-average, 4-good, 5-excellent). The mean ratings for all of the judges were calculated as an indication of the degree of item-competency congruence for each item. The higher the number, the more congruent the item-competency match. The judges were also asked to rate the technical quality of the item by indicating whether they felt the item should be rejected (1), accepted with revisions (2), or accepted as it was (3).

Berk's (1980) guidelines for the selection of items for a criterion-referenced instrument were considered in the decision to retain, revise, or reject each item. The criteria for acceptable items are summarized in the following manner:

Item characteristic	Criterion	Index value
Item-competency congruence	Matches objective measured	mean rating (3.00-5.00)
Difficulty	Difficult for uninstructed group. Easy for instructed group.	0-50 (uninstructed) 50-100 (instructed)

Item characteristic	Criterion	Index value
Discrimination (DIS_{uigd})	Discriminates between instructed and uninstructed criterion groups	positive index (10-100)
Discrimination (item-score correlation)	Discriminates within each criterion group	positive index (.10-1.00)

As a result of the empirical and judgmental item analyses, a total of eighty items were selected for the three sections of the STCL: twenty-four for computer systems, twenty-eight for computer applications, and twenty-three for computer programming. To ensure that the items included in the final test were representative of the domain of computer literacy, items were selected so that each competency was measured by the test. As the test items were being analyzed, it became apparent that some of the competencies needed to be revised to more accurately reflect the actual competencies measured by the test. Competencies that expressed the same basic skills were combined into one competency, and competencies that expressed more than one skill were separated into discrete competencies. The final list of computer literacy competencies contained eighty competencies: twenty-five in computer systems, twenty-five in computer applications, twenty-one in computer programming, and nine in computer attitudes.

The completed version of the STCL was divided into five sections: Background Information, Computer Attitudes (CAIN and BACS), Computer Systems, Computer Applications, and Computer Programming. The instrument was designed so that it could be administered as a whole or by section.

The STCL was administered to two groups of subjects, instructed and uninstructed students, in order to determine performance standards, demonstrate the test's validity and reliability, and to determine normative data. The CAIN was not administered at this time because validity, reliability, and normative data had been collected by Maurer in 1983.

The instructed group of subjects consisted of 152 college students who had received instruction consistent with the computer literacy competencies in a course titled Introduction to Computer Applications. The uninstructed group of subjects consisted of 110 college students who had not taken a computer course based on the computer literacy competencies.

Following the administration of the test to the two criterion groups, performance standards, or cut-off scores, for the total STCL and each of its subtests were determined by the method proposed by Berk (1976). This method, which is similar to the Contrasting-Groups Method (Hambleton, 1980), involved the identification of the point of intersection of the frequency distribution curves for the uninstructed and instructed groups.

The test score corresponding to the point of intersection for each section of the test was then evaluated by the principal researcher and the steering committee to determine whether it was a reasonable and appropriate criterion for classifying students as masters and nonmasters of computer literacy. If the performance standard identified with this method was not considered appropriate, a standard was set based on the judgment of the steering committee.

Internal consistency reliability coefficients for the total STCL, each section, and each criterion group were estimated using the Kuder-Richardson 20 reliability formula (Ebel, 1972). Internal consistency reliability for the BACS was estimated with coefficient alpha (Cronbach, 1970). Since the instrument has characteristics of both norm-referenced and criterion-referenced tests, a reliability of greater than 0.70 was considered an acceptable reliability coefficient (Iowa State University Test and Evaluation Services, 1983).

Construct validity for the total STCL and each of its sections was determined by calculating the amount of difference between the mean score of the instructed and the uninstructed groups. The means for the instructed group should

be significantly higher than the means for the uninstructed group. A statistically significant (p 0.05) difference between the two means was considered necessary to demonstrate that the difference in the means did not occur by chance. A T-test was used to determine if the differences between the means of the two groups were significant at the 0.50 level (Mason and Bramble, 1987).

Decision validity for the entire STCL and each section was determined by summing the percentage of uninstructed students who were classified as nonmasters of computer literacy and the percentage of instructed students who were classified as masters of computer literacy. Students who scored above the cut-off score were classified as masters of the content, and students who scored below the cut-off score were classified as nonmasters of the content. The number of instructed students classified as masters and the number of uninstructed students classified as nonmasters should be high to demonstrate that the decisions made based on the test scores are valid decisions. In other words, the higher the total percentage of students who performed on the test as they were expected to perform, the more decision validity the test contained (Hambleton, 1980).

Normative data for the two criterion groups, uninstructed and instructed college students, were collected for the entire STCL and each section of the test. The mean score, standard deviation, range, frequency distribution curves, and percentile scores were reported for the total STCL and each of its sections.

The final step in the development of the STCL was the preparation of a complete STCL package that was sponsored by the Iowa State University Research Foundation (ISURF) for publication. The components of the complete STCL package are the following: the Achievement Test containing 80 multiple choice test items divided into three sections, Computer Systems, Computer Applications, and Computer Programming; the Computer Anxiety Index containing 26 items; and the Test Administrator's Manual containing the definition of computer literacy, the competency list, the normative group tables, the administration instructions, and the Computer

Literacy Bibliography. The complete STCL package or any of its components may be purchased through the College of Education at Iowa State University. Scoring of the STCL may be accomplished at Iowa State University's Test and Evaluation Service. Individual and group scores, and item analyses are provided for the STCL and its subtests.

Results

The means, standard deviations, and ranges for the STCL and its sections for each criterion group are reported in Table 1. The mean scores for the instructed group were higher than the mean scores for the uninstructed group by one to two standard deviations for all sections of the test except the BACS (Ellsworth and Bowman, 1982). The mean scores for the BACS suggested that both instructed and uninstructed subjects had more positive attitudes about computers than negative attitudes since lower scores indicated positive attitudes toward computers. On a scale of 1-6, with 1 being the most positive, the mean scores coverted to average ratings were 2.27 for the instructed group and 2.50 for the uninstructed group.

The percentile scores and corresponding raw scores for the total STCL and its subtests for each criterion group are presented in Tables 2 and 3. These tables can be used to compare students' scores on subsequent testings of the STCL with similar or differing norm groups. The change in percentile scores for similar norm groups over many subsequent testings could be used to substantiate that students in a particular group, as a whole, are becoming more or less computer literate. The percentile scores of students could be used as guide for placing students in instructional treatments based on whether they score above or below a specified percentile.

The frequency distributions of raw scores for the total STCL and each subtest for both criterion groups are shown in Figures 1 through 5. The differences

in the frequency distributions between the two criterion groups can be more easily seen and evaluated if the two distributions are shown on the same graph. The points of intersection of the two distribution curves for each section of the test were identified as the cut-off scores for classifying students as masters and nonmasters of the content measured by each section of the instrument. The resulting cut-off scores are reported in Table 4.

The steering committee evaluated the cut-off scores established with this method and concluded that the cut-off scores were not appropriate standards for classifying students as masters and nonmasters of computer literacy. The steering committee felt that computer literate individuals should be able to answer more of the items correctly than the cut-off scores suggested. The mean score for the instructed group in each section was considered a more appropriate cut-off score for each section of the test. The cut-off scores recommended by the steering committee are reported in Table 5.

Item statistics for each STCL test item were collected for both criterion groups and are reported in Table 6. The Difficulty Indices were determined by computing the percentage of students in each group who answered the item correctly. A range of 0 to 100 was possible, with 100 indicating the lowest level of difficulty. Two different Discrimination Indices were computed for each item.

The first Discrimination Index computed was the correlation between the item score and the total test score. A range of -1.00 to +1.00 was possible with +1.00 indicating the highest level of discrimination between the students within each group. The second Discrimination Index was the Uninstructed-Instructed Group Difference Discrimination Index (DIS_{uigd}) (Berk, 1980). It was computed by subtracting the Difficulty Index for the uninstructed group from the Difficulty Index for the instructed group, yielding an index of -100 to +100. The higher the index, the more the item discriminated between students in the instructed group and students in the uninstructed group.

Also included in Table 6 are the competencies measured by each item, the item-competency congruence rating from the judgmental analysis, and the number of subjects who omitted each item. An item-competency rating was not reported for items that were written following the judgmental analysis. However, they were written based on the suggestions of the computer specialists and were thus considered to be congruent with the competencies.

The STCL as a whole was shown to be very reliable, as reported in Table 7. The internal consistency reliability estimates were 0.86 for the instructed group and 0.91 for the uninstructed group. The internal consistency reliability estimates for the individual subtests were somewhat lower, but still indicated that the subtests were reliable for at least one of the criterion groups.

Content validity for the test items was established during the development of the instrument. The use of computer specialists to identify competencies that were representative of the domain of content, the writing of test items matched to the competencies, and the item and judgmental analyses all contributed positively to the content validity of the test.

Construct validity was demonstrated for the total STCL and its subtests by calculating the differences between the mean scores of the instructed and uninstructed groups and showing that the differences were statistically significant. The actual differences between the two criterion groups revealed that the mean of the instructed group was at least one standard deviation above the mean of the uninstructed group for every section except the BACS. T-tests were calculated to determine whether the differences were significant. The results of the T-tests, reported in Table 8, indicate that there were significant differences between the mean scores of the two criterion groups for the total STCL and each of its sections. This indicates that the differences did not occur by chance and helps to establish that the instrument is construct valid.

The STCL and each of its subtests were shown to have a high degree of decision validity because a high percentage of the students scored as their criterion group indicated they should. In other words, the higher the percentage of instructed students who scored above the established cut-off score and the higher the percentage of uninstructed students who scored below the cut-off score, the higher the decision validity of the instrument. The established cut-off scores and corresponding percentages are reported in Tables 4 and 5 for each of the methods used for establishing the cut-off scores. Both methods of establishing the cut-off scores yielded percentages of appropriately classified subjects that were high enough to indicate a high degree of decision validity.

Possible Uses of the STCL

There are several possible uses for the STCL. The primary goals for the STCL were to assess the computer literacy of students at various educational levels, and to guide the development of appropriate instruction so that all students could become computer literate. These primary goals could be accomplished by using the STCL in the following ways.

Cut-off scores could be established in order to identify various degrees of computer literacy. Student scores could be compared to these cut-off scores, and could then be used as one of the criteria for assigning students to courses appropriate to their degree of computer literacy. A cut-off score could be identified as one of the criteria necessary for students to test out of a particular course or segment of a course. For example, a student who scored above the cut-off scores on the Computer Systems and the Computer Applications sections of the STCL, but below the cut-off score on the Computer Programming section could be placed in a course dealing strictly with computer programming.

The STCL could be used to identify the specific computer literacy competencies possessed and not possessed by individuals in a particular course or group. The test items on the STCL are keyed to the competencies, so an examination of individual responses to each item could provide appropriate diagnostic information. Instruction could then be designed, or revised, to ensure that the students master the competencies that were lacking.

The STCL could be used to idenfify persons who are highly computer anxious. Anxiety towards computers may inhibit the subjects' interaction with computers and subsequently influence their level of computer literacy. Persons who are highly computer anxious may benefit from treatments designed to decrease their anxiety prior to receiving instruction about the cognitive aspects of computer literacy.

The STCL could be administered to a variety of norm groups for the collection of more comprehensive norm data. The normative information from a wide variety of groups could be used to identify groups in need of instruction. Subsequent administrations of the test to the same groups following instruction could provide valuable information about the quality of the instruction and could influence the revision of the instruction.

Summary of Results

1. The STCL was found to be a valid measure of computer literacy.
2. The STCL was found to be a reliable measure of computer literacy.
3. The STCL could be used as one of the criteria to classify students as computer literate or not computer literate.
4. The STCL could be used to facilitate decisions about the placement of students in computer literacy courses.
5. The STCL could be used to document the computer literacy competencies of students, and could be used in identifying instructional needs and goals.
6. The STCL could be used as a basis for the design and/or revision of appropriate computer literacy instructional programs.

Table 1. Means, standard deviations, and ranges of total STCL and subtests by criterion group

	Section II[a] Attitudes	Section III Systems	Section IV Applications	Section V Programming	Total[b] test
Mean score					
Instructed[c]	38.70	18.62	17.50	11.09	47.20
Uninstructed[d]	42.50	11.70	10.82	5.72	28.24
Standard deviation					
Instructed	9.49	3.83	4.31	3.79	10.18
Uninstructed	8.19	5.11	5.10	4.59	13.15
Actual range					
Instructed	25-64	[illegible]	0-26	0-20	13-69
Uninstructed	23-62	[illegible]	0-23	0-19	2-63
Possible range	17-102	0-29	0-28	0-23	0-80

[a]Part One, the Beliefs About Computers Scale (Ellsworth and Bowman, 1982): the lower the score, the more positive the attitude toward computers.

[b]Total score for Sections III, IV, and V combined.

[c]N=152.

[d]N=110.

BEST COPY AVAILABLE

Table 2. Percentile scores and corresponding raw scores for BTCL and subtests for instructed criterion group

	Raw Scores by Section					
	II[a]		III[b]	IV[c]	V[d]	Total[e]
%tile	Part 1	Part 2				test
100	64	-	27	26	20	69
99	63	5.0	26	25	19	68
98	59	-	-	-	18	-
97	56	4.6	25	-	-	64
95	53	-	24	24	-	62
94	-	4.0	-	-	17	-
93	-	-	-	23	-	61
92	-	3.8	-	-	-	-
91	49	3.6	23	-	16	-
90	-	-	-	-	-	60
89	-	-	-	22	-	-
88	48	-	-	-	-	-
87	-	-	-	-	-	59
86	-	3.4	-	-	15	-
85	-	-	-	-	-	58
84	47	-	22	21	-	-
82	-	-	-	-	-	57
81	46	-	-	-	14	56
78	45	3.2	-	-	-	55
77	-	-	-	20	-	-
76	-	-	-	-	-	54
75	44	-	21	-	-	-
72	43	-	-	-	13	53
70	42	-	-	-	-	-
69	-	3.0	-	-	-	-
68	41	-	-	-	-	-
66	-	-	-	-	-	52
64	-	-	-	19	-	-
63	-	-	20	-	12	51
61	40	-	-	-	-	-
59	-	2.8	-	-	-	50
58	-	-	19	-	-	-
57	-	-	-	-	11	-
56	39	-	-	-	-	49
55	-	-	-	18	-	-
53	-	-	-	-	-	48
50	-	-	-	-	-	47
49	38	-	18	17	-	-
47	-	2.6	-	-	-	-
46	-	-	-	-	10	46
43	-	-	-	-	-	45
41	37	-	-	-	-	-
40	-	-	17	16	-	-
38	36	-	-	-	-	44
35	-	2.4	-	-	-	43
34	35	-	-	-	9	-
31	-	-	-	-	-	42
30	34	-	-	-	-	-
28	-	-	16	15	8	41
24	33	2.2	-	-	-	40
22	-	-	15	-	-	39
14	31	2.0	14	13	-	-
12	30	-	-	12	-	34
9	29	-	-	-	-	33
8	-	1.8	-	-	-	-
4	26	1.6	-	-	-	29
3	25	-	11	9	-	27
2	5	-	10	-	4	-
1	2	1.4	9	7	3	26

[a]Computer Attitudes. Part 1 is the Beliefs About Computers Scale (Ellsworth and Bowman, 1982) with a possible range of 17-102, 17=most positive computer attitude. Part 2 is the Computer Anxiety Index (Maurer, 1983) with a possible range of 1-6, 6=highest level of computer anxiety.

[b]Computer Systems (maximum possible score=19).

[c]Computer Applications (maximum possible score=28).

[d]Computer Programming (maximum possible score=23).

[e]Sections III, IV, V combined (maximum possible score=80).

BEST COPY AVAILABLE

Table 3. Percentile scores and corresponding raw scores for STCU and subtests for uninstructed criterion group

	Raw Scores by Section					
	II[a]					
%tile	Part 1	Part 2	III[b]	IV[c]	V[d]	Total[e] test
100	62	-	25	23	19	63
99	61	5.0	24	-	18	-
98	-	-	-	-	17	62
97	[illegible]	4.6	23	22	-	58
96	57	-	21	-	16	57
95	56	-	20	21	15	55
94	-	4.0	-	-	[illegible]	-
93	-	-	19	19	-	52
92	54	3.8	-	18	12	-
91	53	3.6	-	-	-	48
90	52	-	-	-	-	45
89	-	-	18	17	-	-
88	-	-	17	-	-	44
86	-	3.4	-	-	11	43
85	-	-	-	16	10	42
83	50	-	-	-	-	39
81	-	-	-	-	9	38
80	49	-	16	-	-	37
79	-	-	-	14	-	36
78	48	3.2	-	-	-	-
77	-	-	15	-	-	35
75	-	-	-	13	8	-
74	-	-	-	-	-	34
69	-	3.0	-	-	7	-
68	46	-	13	-	-	33
63	-	-	-	-	-	31
60	-	-	-	-	-	30
59	44	2.8	-	11	-	29
56	-	-	-	-	5	-
51	43	-	-	-	-	27
50	-	-	-	10	-	-
47	-	2.6	11	-	-	-
46	42	-	-	-	-	-
45	-	-	-	9	4	25
43	-	-	10	-	-	-
42	41	-	-	-	-	24
39	40	-	-	-	-	23
36	-	-	-	-	3	-
35	39	2.4	9	8	-	-
31	38	-	-	-	-	-
28	-	-	-	-	2	20
26	36	-	8	-	-	-
25	-	-	-	7	-	19
24	-	2.2	-	-	-	-
23	-	-	-	-	1	-
20	-	-	7	-	-	18
15	-	-	6	6	-	17
14	33	2.0	-	-	-	-
13	-	-	-	5	0	14
10	31	-	5	-	-	12
8	-	1.8	-	4	-	10
7	-	-	4	-	-	9
5	28	-	3	2	-	8
4	-	1.6	2	-	-	5
3	27	-	-	1	-	3
2	25	-	1	0	-	2
1	23	1.4	0	-	-	-

[a]Computer Attitudes. Part 1 is the Beliefs About Computers Scale (Ellsworth and Bowman, 1982) with a possible range of 17-102, 17=most positive computer attitude. Part 2 is the Computer Anxiety Index (Maurer, 1983) with a possible range of 1-6, 6=highest level of computer anxiety.

[b]Computer Systems (maximum possible score=29).

[c]Computer Applications (maximum possible score=28).

[d]Computer Programming (maximum possible score=23).

[e]Sections III, IV, and V combined (maximum possible score=80).

BEST COPY AVAILABLE

Table 4. Cut-off scores (Berk, 1976) for total STCL and subtests and percentages of criterion groups classified as masters and nonmasters of the content

Section of test	Cut-off[a] score	Percentage of criterion groups appropriately classified as masters and nonmasters of the content		
		Uninstructed[b] (N=110)	Instructed[c] (N=152)	Both groups[d] (N=262)
III. Systems	14	68	99	81
IV. Applications	13	70	88	81
V. Programming	6	56	95	79
Total test	37	78	84	81

[a]Cut-off was determined by the point of intersection between the frequency distribution curves for the uninstructed and instructed criterion groups (Berk, 1976).

[b]Percentage of uninstructed group who scored below the cut-off score.

[c]Percentage of instructed group who scored above the cut-off score.

[d]Percentage of both groups who scored above or below the cut-off score as their criterion group indicated they should.

Table 5. Cut-off scores set by steering committee for total STCL and subtests and percentages of criterion groups classified as masters and nonmasters of the content

Section of test	Cut-off[a] score	Percentage of criterion group appropriately classified as masters and nonmasters of the content		
		Uninstructed[b] (N=110)	Instructed[c] (N=152)	Both groups[d] (N=262)
III. Systems	19	90	51	67
IV. Applications	18	90	51	67
V. Programming	11	85	54	67
Total test	47	92	54	70

[a]Cut-off was determined by the judgments of the steering committee to be the mean score for the instructed group.

[b]Percentage of uninstructed group who scored below the cut-off score.

[c]Percentage of instructed group who scored above the cut-off score.

[d]Percentage of both groups who scored above or below the cut-off score as their criterion group indicated they should.

BEST COPY AVAILABLE

Table 1. ATCL test-item analysis data from instructed and uninstructed criterion groups

Item	Comp[a]	I-C[b]	Uninstructed Omits[c]	Uninstructed Diff[d]	Uninstructed Disc[e]	Instructed Omits	Instructed Diff	Instructed Disc	DIS[f] wtgd
1	S1	4.6	19	53	.41	2	69	.32	16
2	S2	4.2	6	80	.21	2	91	.21	11
3	S3	4.0	7	74	.49	0	87	.33	13
4	S4	4.2	[illegible]	55	.36	0	70	.02	15
5	S4	4.6	[illegible]	64	.46	0	78	.22	14
6	S5	4.2	34	25	.27	2	41	.26	16
7	S6	4.6	22	36	.53	0	36	.33	0
8	S7	4.4	13	55	.62	0	82	.34	27
9	S8	-	8	63	.37	0	78	.15	15
10	S9	3.8	20	62	.60	0	89	.10	27
11	S9	4.2	15	53	.48	1	75	.32	22
12	S10	3.7	12	64	.37	0	56	.28	-8
13	S11	4.7	27	20	.20	1	20	.09	0
14	S12	4.6	24	42	.12	0	50	.16	8
15	S13	-	11	39	.23	0	62	.27	23
16	S14	-	11	30	.48	0	39	.18	9
17	S15	4.2	15	61	.63	1	77	.[illegible]	16
18	S16	4.4	4	55	.38	0	70	.41	15
19	S16	4.6	11	46	.39	2	77	.18	31
20	S17	4.4	38	24	.49	[illegible]	55	.25	31
21	S17	4.0	34	37	.54	0	61	.29	24
22	S18	3.0	17	55	.34	1	51	.09	4
23	S19	4.5	12	37	.58	0	43	.25	6
24	S20	-	31	16	.09	1	55	.51	39
	S21	-							
25	S2[illegible]	-	21	47	.49	1	72	.42	25
	S[illegible]	-							
26	S[illegible]	4.8	15	46	.38	0	93	.31	47
27	S23	3.4	20	39	.47	0	63	.30	24
28	S24	-	9	13	.25	0	37	.17	18
29	S25	4.0	7	48	.30	0	92	.19	44
30	A1	3.0	11	47	.04	3	51	.20	3
31	A2	-	11	60	.48	2	55	.44	-5
32	A3	-	9	41	.29	1	64	.30	23
33	A4	4.0	4	87	.47	1	98	.28	11
34	A5	3.0	5	50	.40	3	76	.44	26
	A6	4.5							
35	A7	4.3	5	62	.57	1	87	.38	25
36	A7	4.0	5	21	.23	1	32	.23	11
37	A8	3.8	5	59	.41	1	75	.46	16
38	A9	4.8	9	71	.58	2	93	.25	22
39	A10	4.2	22	52	.60	3	81	.42	29
	A11	-							

[a]Competency that item measures. Code refers to the section of the test (System-S, Applications-A, Programming-P) and the number of the competency from the list of competencies in Appendix E.

[b]Item-competency congruence (mean rating of judges on a scale of 1-5, 5=high congruence).

[c]Number of students who omitted the item.

[d]Difficulty Index (percent of students who answered item correctly, 0-100).

[e]Discrimination Index (item-score correlation, -1.00-1.00).

[f]Uninstructed-Instructed Group Difference Discrimination Index ($Diff_I - Diff_U$, -100 - +100).

BEST COPY AVAILABLE

Table 6. (continued)

Item	Comp[a]	I-C[b]	Uninstructed			Instructed			DIS[f] wtd
			Omits[c]	Diff[d]	Disc[e]	Omits	Diff	Disc	
40	A10	4.6	28	52	.77	2	81	.48	39
	A11	4.2							
41	A12	4.2	21	12	.32	1	82	.21	70
	A13	-							
42	A12	4.8	34	36	.36	2	89	.38	53
	A13	-							
43	A14	-	13	34	.59	1	30	.43	-4
	A15	-							
44	A14	-	12	54	.40	1	33	.30	-21
45	A16	4.3	12	38	.41	1	40	.36	2
	A18	-							
46	A16	-	15	75	.62	1	92	.36	17
	A17	3.6							
	A19	4.4							
47	A18	5.0	11	60	.43	1	62	.42	2
48	A20	4.0	15	17	.09	1	68	.37	49
49	A21	3.6	28	30	.56	1	58	.22	28
50	A22	-	20	59	.66	1	81	.37	25
51	A23	-	13	46	.28	1	48	.27	2
52	A22	-	18	17	.07	2	29	.11	12
	A23	-							
	A25	-							
53	A24	5.0	12	41	.26	1	66	.23	25
	A25	3.8							
54	A24	4.6	16	47	.40	1	70	.42	33
	A25	4.0							
55	A24	-	21	31	.39	2	47	.40	16
	A25	-							
56	A24	-	14	13	.27	1	23	.15	10
	A25	-							
57	A25	-	24	22	.62	1	48	.40	26
58	P1	3.4	17	41	.37	1	44	.24	9
	P2	3.4							
59	P3	4.7	26	45	.75	2	67	.51	22
60	P3	4.6	30	25	.51	2	45	.26	20
61	P4	5.0	25	61	.48	3	64	.21	3
62	P5	5.0	25	71	.58	1	82	.37	11
63	P5	5.0	26	61	.73	1	73	.36	12
64	P6	4.6	34	36	.73	1	50	.50	14
65	P4	4.4	39	46	.66	7	42	.25	-4
	P7	-							
66	P8	3.5	37	15	.46	1	48	.32	33
67	P8	3.5	37	45	.63	1	63	.21	18
68	P9	5.0	39	27	.39	1	45	.22	18
69	P9	5.0	32	49	.67	1	56	.35	7
70	P10	3.8	28	16	.36	1	77	.52	61
71	P11	3.0	34	20	.41	1	25	.24	5
72	P2	4.4	39	21	.56	2	26	.30	5
	P7	-							
	P12	-							
	P14	-							
73	P15	4.6	41	38	.77	2	65	.25	27
74	P15	4.8	41	61	.89	2	80	.33	19
	P16	4.8							
75	P15	4.8	44	30	.58	2	30	.40	0
76	P15	4.0	42	41	72	2	32	.39	-9
	P13	3.8							
77	P17	4.8	42	34	.46	2	31	.15	-3
78	P18	-	72	21	-.07	14	24	.04	3
79	P19	-	72	29	.68	15	37	.33	8
80	P20	-	74	22	.14	21	24	.19	4

BEST COPY AVAILABLE

Table 7. Reliability estimates for total STCL and subtests by criterion group

Criterion group	Section II[a] Attitudes	Section III Systems	Section IV Applications	Section V Programming	Total[b] test
Instructed	0.81[c]	0.64[d]	0.75	0.69	0.86
Uninstructed	0.73	0.78	0.80	0.81	0.91

[a]Part One, the Beliefs About Computers Scale (Ellsworth and Bowman, 1982).

[b]Sections III, IV, and V combined.

[c]Coefficient alpha.

[d]Kuder-Richardson 20 reliability coefficient (Sections III, IV, V, and total).

Table 8. Comparison of STCL test scores of instructed versus uninstructed criterion groups

	Number	Mean	S.D.	T-Value	2-Tailed Probability ($p<.001$)
Section II					
Instructed	152	38.7	9.5	3.38	0.001**
Uninstructed	110	42.5	8.2		
Section III					
Instructed	152	18.6	3.8	-11.92	0.001**
Uninstructed	110	11.7	5.1		
Section IV					
Instructed	152	17.4	4.4	-11.23	0.001**
Uninstructed	110	10.8	5.1		
Section V					
Instructed	152	11.2	3.8	-10.16	0.001**
Uninstructed	110	5.7	4.6		
Total test					
Instructed	152	47.2	10.2	-12.58	0.001**
Uninstructed	110	28.2	13.2		

**$p<.01$.

BEST COPY AVAILABLE

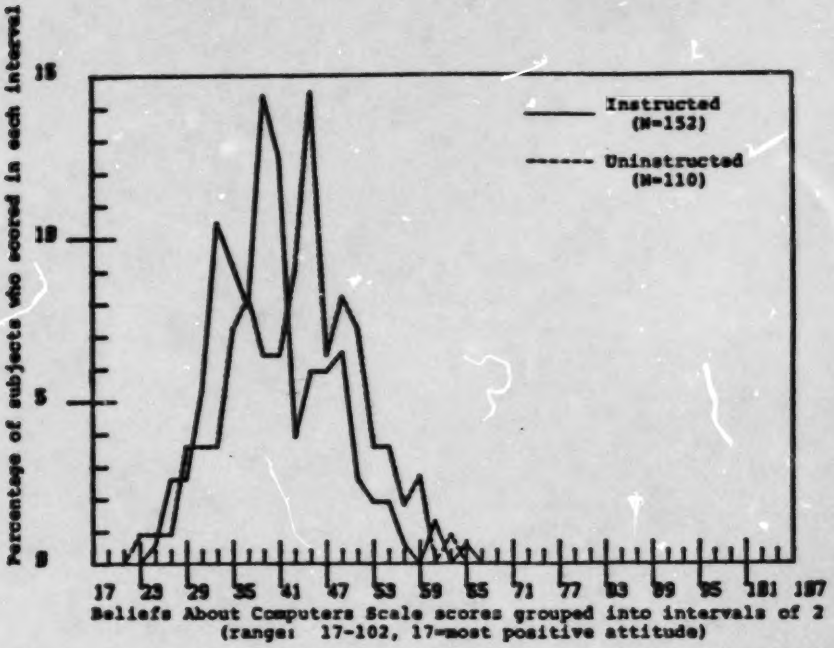

Figure 1. Frequency distributions of scores for WTCL Section II, Part One (Beliefs About Computers Scale) by criterion group

BEST COPY AVAILABLE

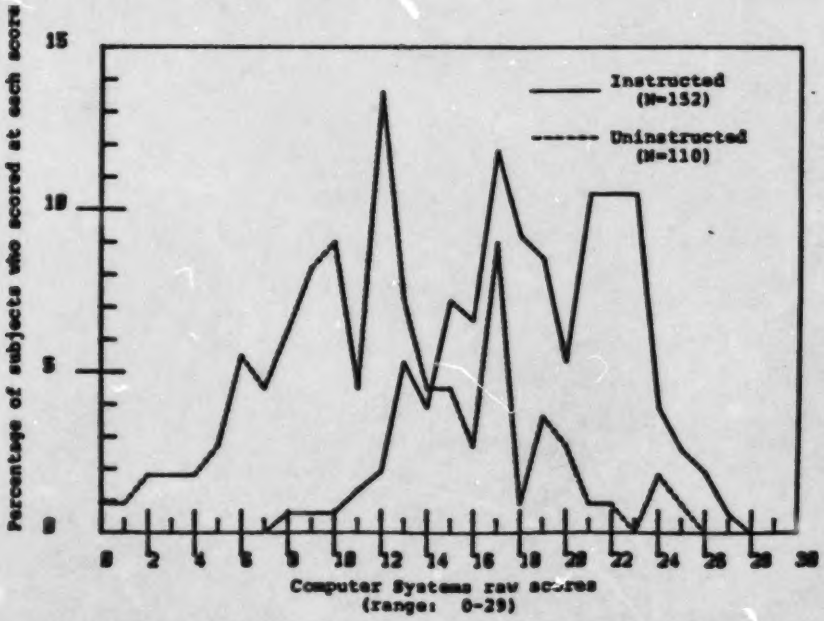

Figure 2. Frequency distributions of scores for WTCL Section III (Computer Systems) by criterion group

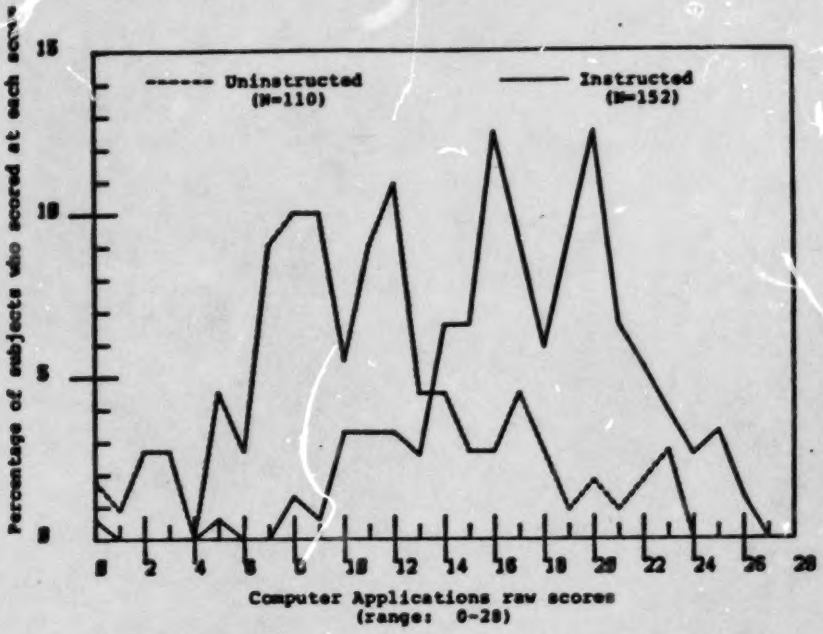

Figure 3. Frequency distributions of scores for STCL Section IV (Computer Applications) by criterion group

BEST COPY AVAILABLE

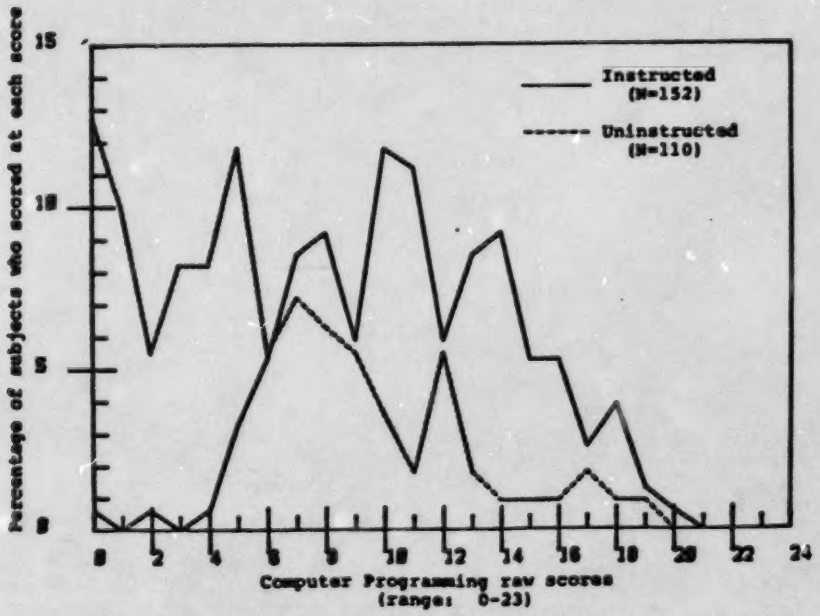

Figure 4. Frequency distributions of scores for STCL Section V (Computer Programming) by criterion group

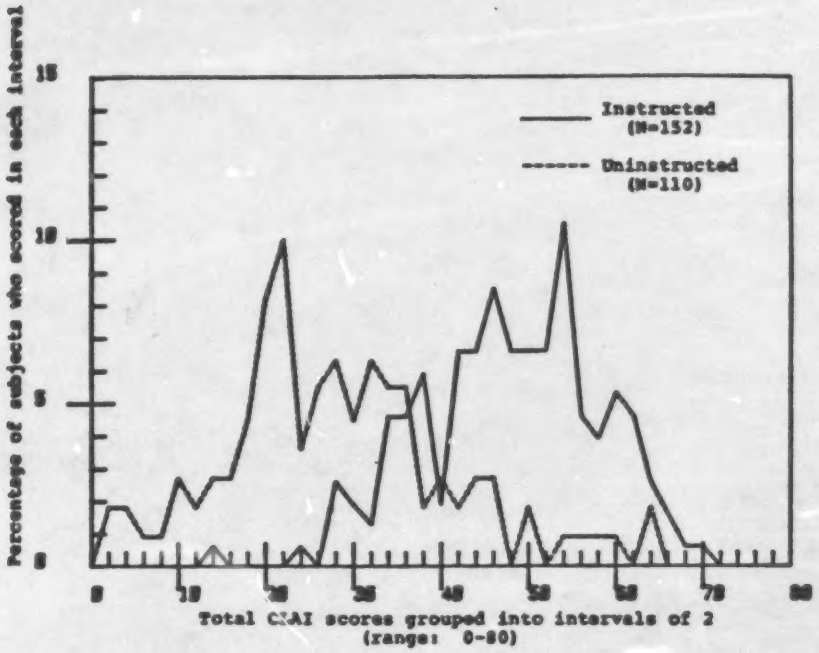

Figure 5. Frequency distributions of scores for total MCL by criterion group

BEST COPY AVAILABLE

References

Berk, R. A. Determincation of optimal cutting scores in criterion-referenced measurement. Journal of Experimental Education, 1976, 45, 4-9.

Berk, R. A. (Ed.). Criterion-referenced measurement: The state of the art. Baltimore: Johns Hopkins University Press, 1980.

Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. Taxonomy of educational objectives: Cognitive domain. New York: David McKay, 1956.

Cronbach, L. J. Essentials of psychological testing (3rd ed.). New York: Harper and Row, 1970.

Ebel, R. L. Essentials of educational measurement. Engelwood Cliffs, New Jersey: Prentice-Hall, Inc., 1972.

Ellsworth, R., & Bowman, B. E. A "beliefs about computers" scale based on Ahl's questionnaire items. The Computing Teacher, December 1982, 10(4), 32-34.

Hambleton, R. K. Test score validity and standard-setting methods. In Berk, R. A. (Ed.), Criterion-referenced measurement: The state of the art. Baltimore: Johns Hopkins University Press, 1980.

Hively, W., Patterson, H. L., & Page, S. H. A "universe-defined" system of arithmetic achievement tests. Journal of Educational Measurement, Winter 1968, 5(4), 275-290.

Iowa State University Test and Evaluation Services. Summary of services: Test and evaluation services. Unpublished manuscript. Iowa State University, 1983.

Kelman, P. (Ed.). The 1983 Classroom Computer News directory of educational computing resources. Watertown, Massachusettes: Intentional Educations, Inc., 1983.

Klassen, D. L., Anderson, R. E., Hansen, T. P., & Johnson, D. C. A study of computer literacy in science education: Final report. 1978-1980. St. Paul: Minnesota Educational Computing Consortium, 1980.

Klein, S. P., & Kosecoff, J. Issues and procedures in the development of criterion referenced tests. 1973. ERIC document ED 083 284.

Mason, E. J., & Bramble, W. J. Understanding and conducting research/Applications in Education and Behavioral Sciences. New York: McGraw-Hill Book Company, 1978.

Maurer, M. M. Development and validation of a measure of computer anxiety. Unpublished Master's Thesis. Iowa State University, 1983.

Millman, J. Criterion-referenced measurement. In Popham, W. J. (Ed.), Evaluation in education: Current applications. Berkely, California: McCutchan Publishing Company, 1974.

BEST COPY AVAILABLE

Moursund, D. International Council for Computers in Education organization members. The Computing Teacher, April 1983, 10(8), 80.

Moursund, D. Summer school course listing: Programming a great summer. The Computing Teacher, March 1982, 9(7), 12-23.

National Commission on Excellence in Education. An open letter to the American people. A nation at risk: The imperative for educational reform, Education Week, April 27, 1983, II(3), 12-16.

Popham, W. J. Domain specification strategies. In Berk, R. A. (Ed.), Criterion-referenced measurement: The state of the art. Baltimore: Johns Hopkins University Press, 1980.

Rovinelli, R. J., & Hambleton, R. K. On the use of content specialists in the assessment of criterion-referenced test item validity. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, 1976.

Seidel, R. J., Anderson, R. E., & Hunter, B. (Ed.). Computer Literacy: Issues and directions for 1985. New York: Academic Press, 1982.

Smith, J., & Moum, G. S. Proceedings of NECC-82 National Educational Computing Conference 1982. Columbia, Missouri: The University of Missouri, 1982.

BEST COPY AVAILABLE